17 September, 1970.

Professor Howard M. Temin
McArdle Laboratory for Cancer Research
University of Wisconsin
Madison, Wisconsin 53706
U.S.A.

Dear Professor Temin

Your letter of September 10 has made me think more about these problems.

First, about RNA→DNA transfer in uninfected cells. Let me say at once that I think people should look hard for these in a fairly large variety of biological situations. This is because, if any cases were found, they would be of considerable importance. What we are discussing, then, is our respective guesses as to how likely these transfers are to be found, and we both agree that they are sufficiently plausible to make a search worth while. The only point at issue, therefore, is are they very probable (as opposed to sufficiently so), and in which context are they likely to be found? The latter point is of some interest, as it would to some extent guide the research.

I agree with your point about long-term storage and stability. I am not so happy about your implicit argument that instability suggests RNA. At first sight, amplification suggests nucleic acid (as you imply also) but here one has to be careful. The ordinary control mechanisms can easily provide amplification in the loose sense. You have to ask what you want to amplify.

Now if you want to provide a few copies of a protein where many are needed the "obvious" mechanism violates the central dogma. An indirect mechanism can work if the nucleic acid which codes the protein sequence is already in the cell concerned. This will normally be the case, and thus in many cases the signal for amplification can be any suitable (small?) molecules and there is no strong reason to invoke either protein or nucleic acid.

The only exceptions I can think of would occur when, for some reason, the coding nucleic acid is not in the relevant cell, or, if it is, it is for some reason difficult to pick it out against a background of similar but different sequences. Either or both of these situations might be invoked in the antibody case. For example, if one believed the somatic mutation theory the

identical sequence information would not be in two cells (unless they were part of the same clone) and thus nucleic acid transfer might be useful, and perhaps RNA→DNA transfer as well.

At the moment I am quite unconvinced that there are situations in either embryology or the nervous system that strongly suggest any requirement of this kind. It seems to me that the obvious (indirect) control mechanisms are quite adequate to cope with what we know. Antibody formation requires a very special mechanism, because the system must respond to an enormous variety of unknown stimulae, and cannot use geographical location to do this. I would argue that there is no requirement corresponding to this in embryology, and that in memory, etc., the main trick used is geographical. That is, weak connections are made, at least to some extent, "at random", and then made stronger by use. I don't think there is a different protein (or combination of proteins) at each and every synapse, although naturally synapses are not all identical, and will fall into classes. Curiously enough, it is instinct, not memory, which presents the more difficult problem.

About unknown transfers, I think we had better agree to differ. I regard them as highly unlikely. Of course I could be wrong, so it is important to have at least some people who don't share my views, as we certainly don't want to miss them if they really exist.

Now about the sequence hypothesis. First I should remark that when I originally invented the name I quite inadvertently defined it wrongly. My definition implied that all DNA codes for protein. I didn't believe this at the time and it was simply due to careless drafting. Curiously enough I appear to be the only person to have noticed this slip.

What it should say is that all protein sequences have been derived from some nucleic acid sequence. However, even here one has to be careful. To begin with, one has to allow for amino acid modification after polypeptide chain synthesis. This would include phosphoserine, hydroxylysine, the change from trypsinogen to trypsin, etc. However, these seem to present little difficulty, and in fact were faced quite early on. The second reservation is more difficult. It implies that the machinery for protein synthesis exists, and that the cell is in a "normal" state as far as temperature, pH, etc. is concerned. At this point one has to bring in the idea of errors, (see Al Hershey's remarks, top of page 699, Nature (1970) 226) and distinguish these from other error-free versions of the machinery. This leads one to the concept of self-consistent machines. If one arbitrarily changed a bit of the code, in an error-free way, would the machinery work if the new version of the code were used to decode the (old) instructions for making the machinery. You can see that all this is

getting rather highbrow! In fact, Sydney tells me he did not cover this idea in his talk at Woods Hole.

I think, therefore, that when you talk about the possibility that the sequence hypothesis might not be true, you should state rather clearly what you mean. In the naive sense it clearly is true. In the highbrow sense it is very difficult to say exactly what it is. I do not believe that one can say, in any simple sense, that all the information is contained in the DNA, unless one brings in the difficult idea of self-consistency. For example, the mechanism of protein synthesis depends on the activating enzymes being made correctly. This cannot be done unless the genetic code is given, and thus cannot be got out of the DNA sequence unless one already knows the code. Of course you are always entitled to wonder whether all you need is the DNA sequence plus the genetic code.

Thus the real question to ask is, how much extra information is required, in addition to the DNA and the code, to make a particular cell work at a particular time? The subsidiary question is how much of this information is due to the environment, and how much to what is inside it, and also how the material within it was produced at some previous time? Was it coded by the DNA (though indirectly) at some previous time, or was there an infinite regress back in time, not depending entirely on the DNA? For example, the cortex of an egg probably contains (in many cases) essential information for the development of the egg. Was this controlled by the DNA in the oocyte? Or was it due to the cortex in some previous cell, which depended again on the cortex of some previous cell, etc. (as in some of Sonneborn's cases)? It is extremely difficult even to state the problem in a really rigorous way, and the above remarks should only be regarded as a sketch.

A fuller discussion would have to distinguish errors from information. The latter should strictly be kept for cases where there are at least two clear-cut alternatives, both of which can, in some sense, work and produce different results. It also requires the concept of self-consistency. For example, the cortex of an egg may have an essential structure, but it might be that any change in it can only produce a non-viable or sterile organism. It is doubtful if this should be classed as information, rather than machinery. On the other hand, if a change in the cortex produced an altered organism which then reproduced the alteration in the cortex, then this clearly ought to be called information. Unfortunately, one can conceive intermediate cases, and these are the ones which produce semantic problems.

About the last half of your last paragraph. My view would be that the advantage of saving a little bit of DNA would not outweigh the enormous difficulty and cost of violating the central dogma, but, as you realize, I am

hopelessly prejudiced on this issue!

Things are likely to be rather hectic for the next few weeks, but after that I hope to find some quiet time to read carefully the papers you so kindly sent me.

Please excuse the length of this letter. Do drop in next time you are in Cambridge, as discussion is so much easier than prolonged correspondence.

With all good wishes,

Yours sincerely

F. H. C. Crick